

## UJI GOODNESS OF FIT DISTRIBUSI GAMMA TERBOBOTI DENGAN STATISTIK KOLMOGOROV-SMIRNOV UNTUK PARAMETER TERESTIMASI

Radiyatul Mardiyah<sup>1)</sup>, Wayan Somayasa<sup>1)</sup>, Herdi Budiman<sup>1)</sup>, Muhammad Kabil Djafar<sup>1)</sup>  
dan Rahmalia Sahupala<sup>1)</sup>

<sup>1)</sup>Program Studi Matematika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Halu Oleo, Kendari, Indonesia  
E-mail: [radiyatulmardiyah17@gmail.com](mailto:radiyatulmardiyah17@gmail.com)

### ABSTRAK

Distribusi gamma terboboti merupakan versi bobot dari distribusi gamma. Tujuan dari penelitian ini adalah untuk menurunkan model uji *goodness of fit* dengan statistik Kolmogorov-Smirnov untuk distribusi gammaterboboti. Menentukan distribusi limit statistik Kolmogorov-Smirnov dibawah  $H_0$  yaitu diturunkan proses limit dari proses empiris dengan asimtotik. Akan ditunjukkan konsistensi dari statistik uji Kolmogorov-Smirnov secara analitik dan menggunakan simulasi Monte-Carlo. Kuantil-kuantil dari statistik Kolmogorov-Smirnov dihipotesis menggunakan simulasi Monte-Carlo untuk menentukan nilai kritis terhadap pengujian hipotesis  $H_0: F(x) = F_0(x)$  vs  $H_1: F(x) \neq F_0(x)$ , dimana menolak  $H_0$  pada berbagai tingkat signifikansi  $\alpha$  jika  $D_n \geq k_{n;1-\alpha}$ . Hasil perhitungan fungsi *power* dari satatistik uji Kolmogorov-Smirnov, Cramer-Von Mises, dan Anderson-Darling menggunakan distribusi eksponensial dan distribusi Weibull untuk berbagai  $n$  dan  $M$  dengan  $\alpha = 1\%$  dan  $\alpha = 5\%$  adalah 1. Jadi dengan nilai *power* yang maksimal tersebut menunjukkan bahwa ketiga uji mempunyai kemampuan yang tinggi untuk mendeteksi  $H_0$  yang salah. Selanjutnya, dilakukan penerapan uji Kolmogorov-Smirnov untuk mengetahui apakah data waktu (dalam hari) kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen yang digunakan berasal dari populasi yang berdistribusi gamma terboboti. Diperoleh hasil pengujian bahwa data berasal dari distribusi gamma terboboti.

**Kata Kunci:** Distribusi Gamma Terboboti, Estimasi Parameter, Distribusi Limit, Simulasi Monte-Carlo, Fungsi Power, Statistik Kolmogorov-Smirnov

### ABSTRACT

*The weighted gamma distribution is a weighted version of the gamma distribution. The purpose of this study was to reduce the goodness of fit test model with kolmogorov-smirnov statistics for weighted gamma distribution. Determining the statistical limit distribution of kolmogorov-smirnov under the  $H_0$  that is reduced the limit process from the empirical process by asymptotic. Will be shown the consistency of the kolmogorov-smirnov test from kolmogorov-smirnov statistics test which conducted analytically and using monte-carlo simulations. Quantiles from Kolmogorov-Smirnov statistics were approached using Monte-Carlo simulations to determine critical values against hypothesis testing  $H_0: F(x) = F_0(x)$  vs  $H_1: F(x) \neq F_0(x)$ , whereas rejecting  $H_0$  at varying degrees of significance  $\alpha$  if  $D_n \geq k_{n;1-\alpha}$ . The calculation results of the power function from the statistic tests of Kolmogorov-Smirnov, Cramer-Von Mises, and Anderson-Darling used the exponential distribution and weibull distribution for various  $n$  and  $M$  with  $\alpha = 1\%$  and  $\alpha = 5\%$  is 1. So with the maximum power value, it indicate that the three tests have a high ability to detect the wrong  $H_0$ . Furthermore, the statistical practice of the Kolmogorov-Smirnov test was carried out to determine whether the time data (in days) of survival of guinea pigs infected with the virulent tubercle bacillus used came from the populations of distributed Weighted Gamma. obtained the results of testing data is from the weighted gamma distribution.*

**Keywords:** Weighted Gamma Distribution, Parameter Estimation, Limiting Distribution, Monte-Carlo Simulation, Power Function, Kolmogorov-Smirnov Statistics

### 1. Pendahuluan

Distribusi gamma adalah distribusi kontinu dengan dua parameter  $\kappa, \theta$  yang mempengaruhi bentuk distribusi gamma. Hal tersebut menunjukkan bahwa bentuk suatu distribusi tidak hanya ditentukan oleh peubah acak saja tetapi juga oleh parameternya. Pengaruh parameter  $\kappa, \theta$  pada distribusi gamma dapat dilihat dari perubahan yang terjadi pada distribusi gamma. Jika parameter distribusi gamma diberikan suatu nilai atau syarat, maka distribusi

gamma akan membentuk distribusi peluang kontinu yang lain (Warella et al, 2021).

korelasi, dan uji berbasis fungsi karakteristik (pembangkit momen) (Arshad dkk., 2003).

Fungsi bobot merupakan fungsi yang disisipkan non-negatif untuk mendapatkan bentuk distribusi gamma terboboti yang ditampilkan sebagai generalisasi distribusi gamma (Sen et al., 2017). Pembelajaran distribusi terboboti dapat digunakan

untuk pemahaman yang lebih baik tentang distribusi standar, dan memberikan teknik penyebaran distribusi untuk fleksibilitas lebih lanjut agar sesuai dengan superior data. Distribusi terboboti sering terjadi dalam penelitian yang berkaitan dengan bio-kedokteran, reliabilitas, ekosistem dan proses percabangan (Ahmad & Ahmad, 2019).

Distribusi gamma terboboti (*weighted gamma distribution*) merupakan versi bobot dari distribusi gamma. Distribusi gamma terboboti memiliki beberapa submodel distribusi yaitu *weighted eksponensial*, gamma dan eksponensial. Fungsi kepadatan peluang (pdf) dari distribusi *weighted gamma* juga dapat dinyatakan sebagai kombinasi linear dari dua pdf gamma. Distribusi ini mengakomodasi peningkatan dan fungsi tingkat kegagalan dan karenanya memiliki penerapan yang lebih luas dalam analisis keandalan dan kelangsungan hidup (Singla et al., 2016).

Uji goodness of fit merupakan metode yang digunakan untuk mengetahui model distribusi peluang suatu populasi dari mana suatu sampel diambil. Dengan kata lain, uji ini digunakan untuk menguji hipotesis bahwa populasi dari suatu sampel memiliki fungsi distribusi tertentu (Daniel, 1989). Uji goodness of fit yang umum adalah Kolmogorov-Smirnov, Cramér-von Mises, dan Anderson-Darling

Konsep dasar dari uji goodness of fit Kolmogorov-Smirnov adalah dengan membandingkan fungsi distribusi empiris sampel dengan fungsi distribusi dari populasi (Suyanto, 2006).

Misalkan  $F$  adalah fungsi distribusi kumulatif (cdf) yang tidak diketahui untuk populasi  $X$  dan  $F_0$  adalah suatu fungsi distribusi kumulatif (cdf) yang diasumsikan. Hipotesis yang menjadi perhatian dalam goodness of fit adalah:

$H_0: F(x) = F_0(x), \forall x \in \mathfrak{R}$  vs  $H_1: \exists x \in \mathfrak{R}, F(x) \neq F_0(x)$ . dengan kata lain  $H_0 =$  populasi mengikuti distribusi tertentu lawan  $H_1 =$  populasi tidak mengikuti distribusi yang dispesifikasikan pada  $H_0$ . Statistik uji Kolmogorov-Smirnov yang tepat untuk menyelesaikan hipotesis diatas adalah sebagai berikut:

$$D_n = \sqrt{n} \sup_{|x| < \infty} |\hat{F}_n(x) - F_0(x; \theta)|$$

dimana  $\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$ . Uji ini menolak  $H_0$  jika nilai statistik Kolmogorov-Smirnov lebih besar dari titik kritis (Fallo et al., 2013).

Dalam makalah ini akan dibahas tentang uji goodness of fit untuk distribusi gamma terboboti berdasarkan statistik uji Kolmogorov-Smirnov. Terdapat beberapa hal yang akan dibahas yaitu pada bagian 3.1 berisi tentang menentukan distribusi limit statistik Kolmogorov-Smirnov yaitu diturunkan proses limit dari proses empiris dengan asimtotik,

selanjutnya akan ditunjukkan konsistensi dari statistik uji Kolmogorov-Smirnov secara analitik dan menggunakan simulasi Monte-Carlo, dan menghitung nilai kuantil-kuantil dari statistik Kolmogorov-Smirnov dihipotesis menggunakan simulasi Monte-Carlo. Pada bagian 3.2 akan dibahas tentang menghitung power dari uji Kolmogorov-Smirnov, uji Cramer-Von Mises, dan uji Anderson-Darling untuk berbagai macam ukuran sampel yang mengikuti model distribusi di bawah  $H_1$ . Pada bagian 3.3 akan dibahas tentang aplikasi penerapan uji goodness of fit Kolmogorov-Smirnov pada data daya tahan marmut.

## 2. Kajian Pustaka

### 2.1 Distribusi Terboboti

Distribusi terboboti adalah salah satu distribusi untuk pemodelan data statistik dan prediksi yang efisien ketika distribusi standar menghasilkan hasil tidak sesuai. Peristiwa yang dapat dimodelkan dengan distribusi terboboti yaitu ketika pengamatan yang diteliti tidak termasuk sebagai sampel acak dari distribusi asli, hal ini terjadi karena beberapa kejadian tidak dapat diamati atau adanya kerusakan yang disebabkan oleh pengamatan dengan menghasilkan nilai yang berkurang atau sistem pengambilan sampel yang tidak sesuai dengan keadaan asli.

Rao (1965) memperkenalkan dan merumuskan secara umum dengan pemodelan data statistik ketika praktek yang biasa menggunakan distribusi standar ditemukan tidak sesuai.

Misalkan  $X$  adalah variabel acak kontinu non-negatif dengan pdf  $f(x)$ , maka pdf dari variabel acak terboboti  $X_w$  adalah sebagai berikut:

$$f(x) = \frac{w(\alpha x) f(x)}{\mu_w}, x > 0 \quad (1)$$

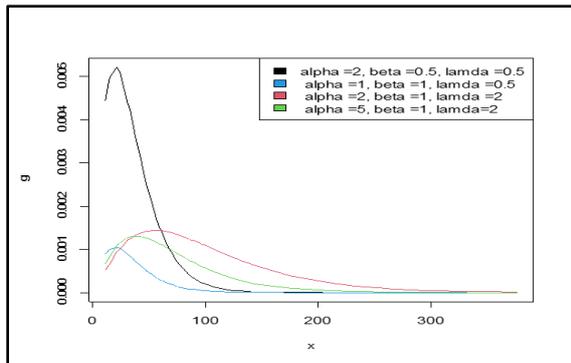
dimana  $w(\alpha x)$  adalah fungsi bobot non-negatif dan  $\mu_w = E[w(X)] < \infty$  (Saghir et al., 2017).

### 2.2 Distribusi Gamma Terboboti

Variabel acak  $X$  dikatakan mengikuti distribusi gamma terboboti (*weighted gamma distribution*) dengan parameter skala  $\lambda$  dan parameter bentuk  $\alpha$  dan  $\beta$ , jika  $X$  mempunyai pdf yang diberikan oleh:

$$f(x; \alpha, \beta, \lambda) = k \frac{(1 - e^{-\alpha \lambda x})^\beta x^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)}, x, \alpha, \beta, \lambda > 0,$$

dimana  $k^{-1} = 1 - \left(\frac{1}{1+\alpha}\right)^\beta$ . Jika  $X$  adalah variabel acak berdistribusi gamma terboboti dengan parameter  $\alpha, \beta$ , dan  $\lambda$ , maka akan digunakan notasi  $X \sim WG(\alpha, \beta, \lambda)$ .



**Gambar 2.2** Grafik pdf distribusi gamma terboboti Fungsi distribusi kumulatif dari  $X \sim WG(\alpha, \beta, \lambda)$  adalah:

$$F(x) = \frac{(1+\alpha)^\beta}{(1+\alpha)^\beta - 1} \left[ G(x; \beta, \lambda) - \frac{1}{(1+\alpha)^\beta} G(x; \beta, \lambda(1+\alpha)) \right]$$

dimana:

$$G(x; \beta, \lambda) = \frac{\lambda^\beta \int_0^x e^{-\lambda t} t^{\beta-1} dt}{\Gamma(\beta)}$$

pdf versi terboboti dari gamma dengan fungsi bobot:

$$w(\alpha x) = 1 - e^{-\alpha \lambda x}, \alpha, \lambda > 0$$

$$\mu_W = \frac{1}{k} = 1 - \left( \frac{1}{1+\alpha} \right)^\beta$$

pada sisi lain,  $\mu_W = E(w(X))$ , sehingga diperoleh:

$$\mu_W = E(w(X))$$

$$= 1 - \left( \frac{1}{1+\alpha} \right)^\beta$$

Distribusi gamma dan eksponensial dapat diperoleh sebagai kasus khusus dari distribusi gamma terboboti untuk nilai parameter tertentu. Kasus khusus tersebut adalah:

a. Jika  $\beta = 1$ , maka diperoleh distribusi eksponensial terboboti dengan pdf:

$$f(x; \alpha, \lambda) = k(1 - e^{-\alpha \lambda x}) \lambda e^{-\lambda x}, x, \alpha, \lambda > 0.$$

b. Jika  $\alpha \rightarrow \infty$ , maka diperoleh:

$$\lim_{\alpha \rightarrow \infty} \frac{1}{1 - \left( \frac{1}{1+\alpha} \right)^\beta} \frac{(1 - e^{-\alpha \lambda x}) \lambda^\beta x^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)} = \frac{\lambda^\beta x^{\beta-1} e^{-\lambda x}}{\Gamma(\beta)}$$

yang merupakan pdf dari distribusi gamma dengan parameter  $\lambda$  dan  $\beta$ .

c. Selanjutnya jika  $\alpha \rightarrow \infty$  dan  $\beta = 1$ , maka berlaku:

$$\lim_{\alpha \rightarrow \infty} \frac{1}{1 - \left( \frac{1}{1+\alpha} \right)^\beta} (1 - e^{-\alpha \lambda x}) \lambda e^{-\lambda x} = \lambda e^{-\lambda x}$$

yang merupakan pdf dari distribusi eksponensial dengan parameter  $\lambda$ .

### 2.3 Estimasi Maksimum Likelihood

Estimasi parameter adalah pendugaan karakteristik populasi (parameter) dengan menggunakan karakteristik sampel (statistik). *Maximum likelihood estimate* atau dikenal juga

sebagaimana metode estimasi maksimum *likelihood* adalah salah satu metode yang paling sering digunakan untuk mencari nilai estimasi parameter dari suatu distribusi (Diantoro, 2013)

Misalkan  $X_1, X_2, \dots, X_n$  merupakan sampel acak dari populasi berdistribusi  $WG(\alpha, \beta, \lambda)$ , maka fungsi *likelihood* dari  $\alpha, \beta$ , dan  $\lambda$  adalah:

$$L(\alpha, \beta, \lambda) = \prod_{i=1}^n \frac{k(1 - e^{-\alpha \lambda x_i}) \lambda^\beta x_i^{\beta-1} e^{-\lambda x_i}}{\Gamma(\beta)}$$

maka Fungsi log *likelihood* untuk  $\alpha, \beta, \lambda$  adalah:

$$\ln L(\alpha, \beta, \lambda) = n \left[ \beta \ln(1 + \alpha) - \ln[(1 + \alpha)^\beta - 1] \right] + \sum_{i=1}^n \ln(1 - e^{-\alpha \lambda x_i}) + n\beta \ln(\lambda) + (\beta - 1) \sum_{i=1}^n \ln(x_i) - \lambda \sum_{i=1}^n x_i - n \ln(\Gamma(\beta))$$

Penduga maksimum *likelihood* untuk  $\alpha, \beta$ , dan  $\lambda$  diperoleh:

$$\frac{\partial \ln L(\alpha, \beta, \lambda)}{\partial \alpha} = 0 \Leftrightarrow \frac{-n\beta(1+\alpha)^{\beta-1}}{(1+\alpha)^\beta((1+\alpha)^\beta - 1)} + \lambda \sum_{i=1}^n \frac{x_i e^{-\alpha \lambda x_i}}{1 - e^{-\alpha \lambda x_i}} = 0 \quad (2)$$

$$\frac{\partial \ln L(\alpha, \beta, \lambda)}{\partial \beta} = 0 \Leftrightarrow n \ln(1 + \alpha) - \frac{n(1 + \alpha)^\beta \ln(1 + \alpha)}{\{(1 + \alpha)^\beta - 1\}} + n \ln \lambda + \sum_{i=1}^n \ln x_i - n\psi(\beta) = 0, \quad (3)$$

$$\frac{\partial \ln L(\alpha, \beta, \lambda)}{\partial \lambda} = 0 \Leftrightarrow \alpha \sum_{i=1}^n \frac{x_i e^{-\alpha \lambda x_i}}{1 - e^{-\alpha \lambda x_i}} + \frac{n\beta}{\lambda} - \sum_{i=1}^n x_i = 0, \quad (4)$$

Karena turunan pertama terhadap  $\alpha, \beta, \lambda$  merupakan fungsi non linear, maka penyelesaian secara konkrit untuk  $\alpha, \beta$ , dan  $\lambda$  tidak dapat dirumuskan, oleh karena itu digunakan pendekatan numeric dengan menerapkan metode iterasi Newton-Raphson. Untuk itu perlu dihitung matriks Hessian yang elemen-elemennya terdiri dari turunan kedua dari fungsi  $\ln L(\alpha, \beta, \lambda)$ .

Sehingga diperoleh turunan parsial kedua dari persamaan (2), (3), dan (4) sebagai berikut:

$$\frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \alpha^2} \Bigg|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} = \frac{n\beta(1+\alpha)^\beta - n\beta + n\beta^2(1+\alpha)^\beta}{(1+\alpha)^2((1+\alpha)^\beta - 1)^2} - \sum_{i=1}^n \frac{x_i^2 \lambda^2 e^{\alpha \lambda x_i}}{(1 - e^{\alpha \lambda x_i})^2}$$

$$\begin{aligned} & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \alpha \partial \beta} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} \\ &= \frac{-n(1+\alpha)((1+\alpha)^\beta - 1 - \beta \ln(1+\alpha)(1+\alpha)^\beta)}{(1+\alpha)^2((1+\alpha)^\beta - 1)^2} \\ & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \alpha \partial \lambda} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} = \sum_{i=1}^n \frac{x_i e^{\alpha \lambda x_i} - x_i - x_i^2 \alpha \lambda e^{\alpha \lambda x_i}}{(e^{\alpha \lambda x_i} - 1)^2} \\ & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \beta \partial \alpha} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} = \frac{n}{(1+\alpha)} \\ & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \beta \partial \lambda} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} = \frac{n \psi'(\beta)}{(1+\alpha)^\beta - 1} \\ & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \lambda \partial \alpha} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} = \sum_{i=1}^n \frac{x_i e^{\alpha \lambda x_i} - x_i - x_i^2 \alpha \lambda e^{\alpha \lambda x_i}}{(1 - e^{\alpha \lambda x_i})^2} \\ & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \lambda \partial \beta} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} = \frac{n}{\lambda} \\ & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \lambda^2} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} = - \sum_{i=1}^n \frac{x_i^2 \alpha^2 e^{\alpha \lambda x_i}}{(1 - e^{\alpha \lambda x_i})^2} - \frac{n\beta}{\lambda^2} \end{aligned}$$

Maka diperoleh matriks Hessian sebagai berikut:

$$H(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = \begin{bmatrix} \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \alpha^2} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \alpha \partial \beta} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \alpha \partial \lambda} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} \\ \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \beta \partial \alpha} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \beta^2} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \beta \partial \lambda} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} \\ \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \lambda \partial \alpha} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \lambda \partial \beta} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} & \left. \frac{\partial^2 \ln L(\alpha, \beta, \lambda)}{\partial \lambda^2} \right|_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}} \end{bmatrix}$$

Sehingga dapat didefinisikan algoritma iterasi metode Newton-Raphson tiga variabel untuk  $\alpha, \beta$  dan  $\lambda$  yaitu:

$$\begin{pmatrix} \alpha_{i+1} \\ \beta_{i+1} \\ \lambda_{i+1} \end{pmatrix} = \begin{pmatrix} \alpha_i \\ \beta_i \\ \lambda_i \end{pmatrix} - \left[ \left( H(\hat{\alpha}_i, \hat{\beta}_i, \hat{\lambda}_i) \right)^{-1} g(\hat{\alpha}_i, \hat{\beta}_i, \hat{\lambda}_i) \right]$$

dimana:

$$g(\hat{\alpha}_i, \hat{\beta}_i, \hat{\lambda}_i) = \begin{bmatrix} \frac{\partial(\ln L(\alpha, \beta, \lambda))}{\partial \alpha} \\ \frac{\partial(\ln L(\alpha, \beta, \lambda))}{\partial \beta} \\ \frac{\partial(\ln L(\alpha, \beta, \lambda))}{\partial \lambda} \end{bmatrix}_{\substack{\alpha=\hat{\alpha} \\ \beta=\hat{\beta} \\ \lambda=\hat{\lambda}}}$$

Selanjutnya akan dilakukan perhitungan iterasi Newton-Raphson pada datayang berdistribusigamma terboboti menggunakan bantuansoftware R,sesuai dengan algoritma metode Newton-Raphson yang berada di sub Bab 2.5.Padasimulasidengan bantuan softwareRini,akandigunakandatabangkitanyangberju mlah 72 data dengan maksimum iterasi100, toleransi error  $\varepsilon = 0,0001$ .

Berdasarkan hasil perhitungan dengan menggunakan bantuan software R, makadiperoleh:

$$\hat{\alpha} = 0,3409,$$

$$\hat{\beta} = 0,000006,$$

$$\hat{\lambda} = 0,00003.$$

## 2.4 Konsep Konvergensi dalam Distribusi

Misalkan  $X_1, \dots, X_n$  merupakan sampel acak dari suatu populasi dengan mean  $\mu$  dan variansi  $\sigma^2$ , dimana  $-\infty < \mu < \infty$  dan  $-\infty < \sigma^2 < \infty$  tetapi model distribusi peluang dari populasinya tidak diketahui, maka eksak dari  $\frac{\sqrt{n}(X-\mu)}{\sigma}$  akan sulit diturunkan.

**Definisi 2.4 (Somayasa, 2019)** Misalkan  $\{X_n\}_{n \geq 1}$  adalah suatu barisan variabel acak dan  $\{G_n\}_{n \geq 1}$  barisan fungsi distribusi yang bersesuaian dengan  $\{X_n\}_{n \geq 1}$ , dimana untuk  $n \geq 1, G_n(y) := \mathbb{P}\{X_n \leq y\}, \forall y \in \mathbb{R}$ . Selanjutnya misalkan  $X$  adalah suatu variabel acak dengan CDF  $F_X$ . Barisan  $\{X_n\}_{n \geq 1}$  dikatakan konvergen dalam distribusi ke  $X$  (dinotasikan dengan  $X_n \xrightarrow{D} X$ ) untuk  $n \rightarrow \infty$ , jika dan hanya jika:

$$\lim_{n \rightarrow \infty} G_n(y) = F_X(y), \forall y \in C(F_X) \quad (7)$$

dimana  $C(F_X)$  adalah himpunan titik-titik yang kontinuitas dari  $F_X$ . Selanjutnya,  $F_X$  disebut fungsi distribusi limit dari  $X_n$ .

### Teorema 2.5 (Teorema Limit Pusat Multivariat) (Serfling, 1992. Hal:29)

Misalkan  $\{X_i\}$  vektor acak berdimensi  $p$  yang saling bebas dan berdistribusi identik dengan mean  $\mu$ , matriks kovarian  $\Sigma$ , maka

$$\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n X_i - \mu \right) \xrightarrow{D} N_p(\mathbf{0}, \Sigma) \quad (8)$$

### Teorema 2.6 (Teorema Pemetaan Kontinu) (Billingsley, 1999. Hal:26)

Misalkan  $(Y_n)_{n \geq 1}$  adalah barisan proses stokastik pada suatu ruang matrik  $S$ , dan  $h := S \rightarrow \mathbb{R}$  adalah pemetaan kontinu pada  $S$  jika  $Y_n \xrightarrow{D} Y$  maka  $h(Y_n) \xrightarrow{D} h(Y)$ .

## 2.5 Gerak Brown

Proses Stokastik gerak brown (*Brownian motion*) pertama kali dikenalkan pada awal abad ke-18 oleh Robert Brown yang mengamati fenomena partikel-partikel kecil yang tersuspensi dalam fluida. Pemodelan secara matematika dari gerak Brown

diperkenalkan oleh Albert Einstein tahun 1905 (Pinsky dan Karlin, 2007).

**Definisi 2.7 (Ross, 1996)** Jika  $\{B(t), t \geq 0\}$  adalah gerak Brown standar, maka  $\{B_0(t), 0 \leq t \leq 1\}$  yang didefinisikan dengan

$$B_0(t) := B(t) - tB(1)$$

disebut sebagai jembatan brown (*Brownian bridge*).

**Teorema 2.8 (Lehmann dan Romano, 2005)**

Misalkan  $B_n(t) = \sqrt{n}(\hat{F}_n(t) - F_0(t))$  merupakan suatu variabel acak bernilai real. Untuk setiap  $t \in \mathbb{R}$ ,  $B_n(t)$  dapat dipandang sebagai fungsi acak pada interval  $[0,1]$ , yang disebut sebagai proses empiris. Dengan menerapkan teorema limit pusat multivariat, jika  $H_0$  benar maka untuk setiap  $t_1, t_2, \dots, t_p$ ,

berlaku  $[B_n(t_1), \dots, B_n(t_p)] \xrightarrow{D} [B_n(t_1), \dots, B_n(t_p)]$ , dimana  $[B_n(t_1), \dots, B_n(t_p)]$  berdistribusi normal multivariat dengan *mean* 0 dan *covariance matrix*  $\Sigma$ . Untuk elemen ke- $(i, j)$ , entri  $\sigma_{i,j}$  bernilai

$$\sigma_{i,j} = \begin{cases} F_0(t_i)(1 - F_0(t_i)) & i = j \\ F_0(\min(t_i, t_j)) - F_0(t_i)F_0(t_j) & i, j \text{ yang lain.} \end{cases}$$

Dengan menerapkan teorema pemetaan kontinu, untuk setiap  $t_1, t_2, \dots, t_p$ , maka

$$\text{Sup}_{1,2,\dots,p} \sqrt{n} |\hat{F}_n(t_i) - F_0(t_i)| \xrightarrow{D} \text{Sup}_{1,2,\dots,p} [B(t_i)].$$

Gerak brown dari  $B(\cdot)$  dapat direpresentasikan sebagai proses kontinu pada interval  $[0,1]$ .

## 2.6 Fungsi Power

Dalam uji hipotesis ruang parameter  $\Theta$  dibagi menjadi dua himpunan bagian yang saling asing, yaitu  $\Theta_0 \subset \Theta$  dan  $\Theta_1 := \Theta - \Theta_0$ . Jika diberikan sampel yang diambil dari populasi  $f_X(x; \theta), \theta \in \Theta$ , prosedur uji hipotesis harus mampu menentukan apakah  $H_0$  ditolak atau diterima (Somayasa, 2019).

**Definisi 2.9(Somayasa, 2019)** Fungsi kuasa (*power function*) dari uji  $\psi$  adalah suatu fungsi  $G_\psi: \theta \rightarrow [0,1]$  yang diberikan oleh

$$G_\psi(\theta) := P\{\text{Tolak } H_0 | \theta \in \Theta\} = E(\psi)$$

untuk  $\theta \in \Theta$ , yang menyatakan peluang menolak  $H_0$  jika  $\theta$  adalah parameter yang sebenarnya. Selanjutnya, ukuran (*size*) dari uji  $\psi$  adalah  $\sup_{\theta \in \Theta_0} G_\psi(\theta)$ . Untuk suatu bilangan  $\alpha \in (0,1)$ , uji  $\psi$  dikatakan uji dengan tingkat signifikansi  $\alpha$ , jika  $G_\psi(\theta) \leq \alpha, \forall \theta \in \Theta_0$ , sehingga jika  $H_0$  merupakan hipotesis sederhana, maka tingkat signifikansi adalah peluang kesalahan tipe I.

## 3. Hasil dan Pembahasan

### 3.1 Distribusi Limit

Misalkan diberikan sampel acak  $X_1, \dots, X_n$  yang diambil dari suatu populasi  $X$  dengan cdf  $F(x), x \in \mathfrak{R}$ . Misalkan  $F$  merupakan cdf yang tidak diketahui untuk populasi  $X$  dan  $F_0$  merupakan fungsi distribusi

kumulatif dari model distribusi gamma berbobot. Hipotesis yang akan menjadi perhatian pada penelitian ini adalah:

$$H_0: F(x) = F_0(x; \alpha, \beta, \lambda), \forall x \in \mathfrak{R} \text{ vs } H_1: \exists x \in \mathfrak{R} F(x) \neq F_0(x; \alpha, \beta, \lambda) \quad (9)$$

Hipotesis (9) dapat diuji dengan menggunakan statistik Kolmogorov-Smirnov, yang didefinisikan sebagai berikut:

$$D_n = \sqrt{n} \sup_{|x| < \infty} |\hat{F}_n(x) - F_0(x; \hat{\alpha}, \hat{\beta}, \hat{\lambda})| \quad (10)$$

Teorema berikut memberikan distribusi limit dari proses empirik

$$\sqrt{n} (\hat{F}_n(x) - F_0(x; \hat{\alpha}, \hat{\beta}, \hat{\lambda})), x \in \mathfrak{R}.$$

**Teorema 3.1** Misalkan  $B_n(t) = \sqrt{n} (\hat{F}_n(t) - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  untuk  $t \in \mathfrak{R}$ , maka berlaku:

$$B_n(t) \xrightarrow{D} B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda})), \text{ untuk } n \rightarrow \infty,$$

dimana  $B^\circ(F_0(t))$  adalah jembatan Brown, yaitu suatu proses stokastik, dengan sifat-sifat berikut:

$$E(B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))) = 0, \quad \forall t \in \mathfrak{R}$$

dan

$$\text{Cov}(B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda})), B^\circ(F_0(s; \hat{\alpha}, \hat{\beta}, \hat{\lambda})))$$

=

$$\min[F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}), F_0(s; \hat{\alpha}, \hat{\beta}, \hat{\lambda})] - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda})F_0(s; \hat{\alpha}, \hat{\beta}, \hat{\lambda}), \text{ untuk } t, s \in \mathfrak{R}.$$

Akan ditunjukkan bahwa untuk sembarang partisi  $t_1 < t_2 < \dots < t_p, p \geq 1$ , vektor

$(B_n(t_1), \dots, B_n(t_p))^T$  konvergen dalam distribusi ke

vektor  $(B^\circ(F_0(t_1)), \dots, B^\circ(F_0(t_p)))^T$ , dimana

$(B^\circ(F_0(t_1)), \dots, B^\circ(F_0(t_p)))^T$  mempunyai

distribusi normal multivariat dengan mean 0 dan matriks kovariansi  $\Sigma = (\sigma_{k,m})_{k,m \geq 1}$ .

Berdasarkan definisi dari  $B_n(t)$ , maka vektor  $(B_n(t_1), \dots, B_n(t_k))$  dapat dituliskan sebagai berikut:

$$B_n := \begin{pmatrix} B_n(t_1) \\ B_n(t_2) \\ \vdots \\ B_n(t_p) \end{pmatrix} = \sqrt{n} \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} (\mathbb{1}\{X_i \leq t_1\} - F_0(t_1)) \\ (\mathbb{1}\{X_i \leq t_2\} - F_0(t_2)) \\ \vdots \\ (\mathbb{1}\{X_i \leq t_p\} - F_0(t_p)) \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{Z}_i$$

dimana  $\mathbf{Z}_i := (Z_{i1}, \dots, Z_{ip})^T$  dengan  $Z_{ik} := \mathbb{1}\{X_i \leq t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}\} - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})$ , untuk  $k = 1, 2, \dots, p$ . Selanjutnya diperoleh untuk sembarang nilai  $k$ :

$$E(Z_{ik}) = E(\mathbb{1}\{X_i \leq t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}\} - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})) = 0$$

sehingga  $E(\mathbf{Z}_i) = \mathbf{0} := (0, 0, \dots, 0)^T \in \mathbb{R}^p$  dan variansi dari  $Z_{ik}$  adalah:

$$\begin{aligned} \text{Var}(Z_{ik}) &= \text{Var}\left(\mathbb{1}\{X_i \leq t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}\} - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})\right) \\ &= F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}) \left(1 - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})\right), \end{aligned}$$

untuk  $k \neq m$ , maka diperoleh

$$\begin{aligned} \text{Cov}(Z_{ik}, Z_{im}) &= \text{Cov}\left(\mathbb{1}\{X_i \leq t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}\} - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}), \mathbb{1}\{X_i \leq t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda}\} - F_0(t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda})\right) \\ &= E\left(\left(\mathbb{1}\{X_i \leq t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}\} - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})\right)\left(\mathbb{1}\{X_i \leq t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda}\} - F_0(t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda})\right)\right) \end{aligned}$$

$$\begin{aligned} &= F_0(\min\{t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}, t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda}\}) \\ &\quad - 2F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})F_0(t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda}) \\ &\quad + F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})F_0(t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda}) \\ &= \min\{F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}), F_0(t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda})\} - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})F_0(t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda}). \end{aligned}$$

jadi  $E(\mathbf{Z}_i) = \mathbf{0}$  dan  $\text{Cov}(\mathbf{Z}_i) = \mathbf{\Sigma} = (\sigma_{k,m})_{k,m \geq 1}$ ,

dimana:

$$\sigma_{k,m} = \begin{cases} F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})(1 - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})), & k = m \\ \min\{F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}), F_0(t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda})\} - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})F_0(t_m; \hat{\alpha}, \hat{\beta}, \hat{\lambda}), & k \neq m \end{cases}$$

dengan menggunakan Teorema Limit Pusat Multivariat diperoleh:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{Z}_i \xrightarrow{D} N_p(\mathbf{0}, \mathbf{\Sigma}), \text{ untuk } n \rightarrow \infty \quad (11)$$

dimana  $N_p(\mathbf{0}, \mathbf{\Sigma})$  adalah distribusi dari vektor  $(B^\circ(F_0(t_1)), \dots, B^\circ(F_0(t_p)))^T$ .

Jadi terbukti bahwa  $\sqrt{n}(\hat{F}_n(t) - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  konvergen dalam distribusi ke  $B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$ ,  $\forall t \in \mathfrak{R}$ .

### 3.1.1 Simulasi Monte Carlo $B_n(t) \xrightarrow{D} B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$

Konvergensi  $B_n(t) \xrightarrow{D} B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  dapat dipergunakan secara visual dengan simulasi Monte Carlo. Adapun algoritma simulasi adalah sebagai berikut:

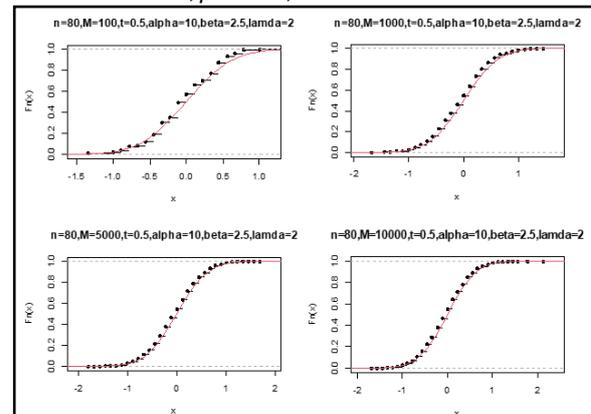
1. Tentukan nilai  $t$ .
2. Bangkitkan  $n$  bilangan acak  $x_1, \dots, x_n$  yang saling bebas dari distribusi unifom  $[0,1]$ .
3. Estimasi parameter  $\alpha, \beta$ , dan  $\lambda$ .
4. Menghitung  $B_n(t) = \sqrt{n}(\hat{F}_n(t) - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  sebanyak  $M$  kali.
5. Membandingkan plot grafik  $ecdf B_n(t)$  dan grafik  $cdf$  dari  $B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  yang berdistribusi normal multivariat dengan mean 0 dan variansi  $F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})(1 - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$ .
6. Ulangi langkah 1-5 untuk berbagai pilihan dari  $n, M$ , dan  $t$ .

Simulasi pertama ditampilkan pada Gambar (3.1) dilakukan dengan menentukan nilai  $t$  yang tetap, kemudian menggunakan data bangkitan  $n = 80$  dan banyaknya perulangan  $M$  yang bervariasi yaitu sebanyak 100, 1000, 5000, dan 10.000 kali.

Kemudian untuk parameter  $\alpha = 10, \beta = 2.5$ , dan  $\lambda = 2$ . Simulasi kedua ditampilkan pada gambar (3.2) dilakukan dengan memilih banyaknya perulangan sebanyak 10.000 kali dan menggunakan data bangkitan  $n$  yang bervariasi sebanyak  $n = 50, n = 70, n = 90$ , dan  $n = 100$ . Kemudian untuk parameter  $\alpha = 10, \beta = 2.5$ , dan  $\lambda = 2$ .

Hasil simulasi yang ditampilkan pada gambar (3.1) dan gambar (3.2) terdiri dari 4 buah grafik yang masing-masing terdapat 2 buah kurva. Kurva berwarna merah adalah grafik jembatan Brown  $B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  yang dibandingkan dengan grafik  $ecdf$  dari  $B_n(t)$  yang ditandai dengan warna hitam.

Kedua grafik memperlihatkan suatu karakteristik, dimana semakin besar nilai  $M$  yang dipilih maka kedua grafik tersebut menunjukkan suatu kecenderungan bahwa  $CDF$  dari  $B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  dapat didekati dengan sangat baik oleh  $ecdf$  dari  $\sqrt{n}(\hat{F}_n(t) - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$ . Sedangkan jika semakin besar nilai  $n$  yang dipilih untuk perulangan  $M = 10.000$  kali, maka kedua grafik tersebut menunjukkan suatu  $CDF$  dari  $B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  dapat didekati dengan sangat baik oleh  $ecdf$  dari  $\sqrt{n}(\hat{F}_n(t) - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  dengan memilih  $\alpha = 10, \beta = 2.5$ , dan  $\lambda = 2$ .

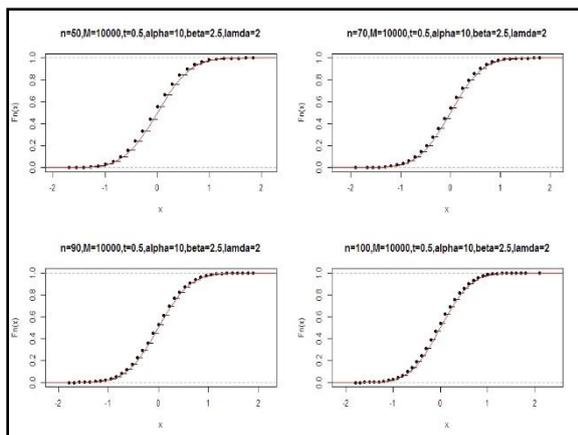


Gambar 3.1. Hasil simulasi untuk  $n = 80, M = 100, 1000, 5000, 10.000, \alpha = 10, \beta = 2.5, \lambda = 2$ .

### 3.1.2 Menghitung Kuantil-Kuantil $D_n$

Misalkan  $k_{n;1-\alpha}$  adalah konstanta sedemikian sehingga  $P\{D_n \leq k_{n;1-\alpha}\} = 1 - \alpha$ . Nilai-nilai dari  $k_{n;1-\alpha}$  untuk berbagai  $n$  dihitung dengan bantuan komputer dengan algoritma berikut:

1. Menentukan taraf signifikansi  $\alpha$ .
2. Membangkitkan  $n$  bilangan acak  $x_1, \dots, x_n$  yang saling bebas dari distribusi uniform  $[0,1]$ .



**Gambar 3.2.** Hasil simulasi untuk  $n = 50, 70, 90, 100, M = 10.000, \alpha = 10, \beta = 2.5, \lambda = 2$

3. Mengitung nilai statistik Kolmogorov-Smirnov.
4. Mengulang langkah 1 – 3 sebanyak  $M$  kali.
5. Menghitung kauntil ke  $\alpha$  dari statistik Kolmogorov-Smirnov.

Pada penelitian ini kuantil-kuantil untuk distribusi peluang statistik uji Kolmogorov-Smirnov dapat dihitung. Akan tetapi, untuk ukuran sampel yang besar perhitungan menjadi sangat rumit. Untuk itu dilakukan pendekatan (aproksimasi) dengan simulasi Monte-Carlo. Perhitungan terhadap kuantil-kuantil dari  $D_n$  yang kemudian akan digunakan untuk mengkonstruksikan daerah penolakan terhadap uji hipotesis (9).

Serfling (1980) dan Khoshnevisan (2006) menyajikan rumus pendekatan  $P\left\{ \sup_{0 \leq x \leq 1} |B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))| \leq x \right\}$  untuk setiap  $x > 0$ , yaitu:

$$P\left\{ \sup_{0 \leq x \leq 1} |B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))| \leq x \right\} \approx 1 - 2 \sum_{m=1}^{\infty} (-1)^{m+1} e^{-2m^2 x^2}, \quad x > 0.$$

Untuk ukuran sampel berhingga  $n$ , baik kuantil-kuantil maupun CDF dari statistik Kolmogorov-Smirnov tidak dapat dihitung secara eksak. Sedangkan kuantil-kuantil dari  $\sup_{0 \leq x \leq 1} |B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))|$  dapat diaproksimasi dengan menggunakan rumus pendekatan.

Dengan menggunakan statisti kuji Kolmogorov-Smirnov, tolak  $H_0$  pada tingkat signifikansi  $\alpha$  jika  $D_n \geq k_{1-\alpha}$ , dimana  $k$  adalah kuantil dari distribusi sampling statistik  $\sup_{0 \leq x \leq 1} |B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))|$ . dimana  $k_{n;1-\alpha}$  adalah konstanta yang memenuhi:

$$P\left\{ \sup_{0 \leq x \leq 1} |B^\circ(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))| \leq k_{1-\alpha} \right\} = 1 - \alpha.$$

Nilai-nilai dari  $k_{n;1-\alpha}$  memberikan hampiran terhadap nilai-nilai dari  $k_{1-\alpha}$ . Penggunaan dari kuantil-kuantil pada Lampiran 6 dalam pengujian terhadap Hipotesis (9) bergantung pada ukuran sampel yang diasumsikan. Sebagai contoh untuk ukuran sampel  $n = 50$  dari tingkat signifikansi  $\alpha = 1\%$ , maka nilai dari  $k_{n;1-\alpha}$  adalah 1.737189.

Oleh karena itu dalam penerapannya,  $H_0$  akan ditolak pada tingkat signifikansi  $\alpha = 1\%$  jika nilai dari statistik  $D_n$  yang dihitung terhadap sampel memenuhi  $D_n > 1.737189$ .

Dari uji terhadap Hipotesis (9) dapat dikonstruksikan daerah kepercayaan untuk fungsi  $F$  yaitu dengan cara memperhatikan daerah penerimaan dari uji tersebut. Jika  $H_0$  ditolak pada tingkat signifikansi  $\alpha$ , maka peluang  $H_0$  tidak ditolak padahal  $H_0$  benar adalah  $1 - \alpha$ . Jadi himpunan semua fungsi  $F$  sedemikian hingga  $D_n \leq k_{n;1-\alpha}$  akan memberikan daerah kepercayaan untuk  $F$ . Sebagai contoh untuk kasus  $n = 50$  dan  $\alpha = 1\%$ , maka himpunan yang didefinisikan sebagai

$$\left\{ F \mid \sup_{|x| < \infty} |\hat{F}_n(x) - F(x)| \leq 1.737189 \right\}$$

merupakan daerah kepercayaan  $100(1 - \alpha)\%$  untuk fungsi  $F$ .

### 3.2 Fungsi Power

Misalkan dilakukan uji goodness of fit untuk populasi berdistribusi gamma terboboti berdasarkan sampel acak  $X_1, X_2, \dots, X_n$ . Fungsi *power* memberikan suatu ukuran apakah suatu uji yang digunakan baik atau tidak. Fungsi *power* dari uji Kolmogorov-Smirnov pada tingkat signifikansi  $\alpha$  diberikan oleh:

$$P\{\text{Tolak } H_0\} = P\{D_n \geq k_{n;1-\alpha}\}.$$

Idealnya uji yang baik sekarang-kurangnya meminimumkan  $P\{\text{kesalahan tipe II}\}$ . Jika  $H_0$  benar, dengan kata lain  $F(x) = F_0(x)$ , maka *power* dari Statistik uji Kolmogorov-Smirnov pada tingkat signifikansi  $\alpha$  adalah

$$\begin{aligned} \text{Power} &= P\{\text{Tolak } H_0 | H_0 \text{ benar}\} \\ &= P\{D_n \geq k_{n;1-\alpha}\} = \alpha \end{aligned}$$

Sebaliknya, jika  $H_1$  benar, yaitu  $F(x) \neq F_0(x)$ , maka *power* dari Statistik uji Kolmogorov-Smirnov pada tingkat signifikansi  $\alpha$  adalah

$$\begin{aligned} \text{Power} &= P\{\text{Tolak } H_0 | H_0 \text{ salah}\} \\ &= P\{D_n \geq k_{n;1-\alpha} | F(x) \neq F_0(x)\} \\ &= 1 - P\{D_n < k_{n;1-\alpha} | F(x) \neq F_0(x)\} \\ &= 1 - \beta. \end{aligned}$$

dimana  $\beta$  adalah peluang kesalahan tipe II. Karena kedua kesalahan dalam uji hipotesis tidak dapat diminimumkan secara bersamaan, jalan yang dapat dilakukan adalah dengan memilih terlebih dahulu bilangan kecil  $\alpha$ , biasanya dipilih  $\alpha = 0,01$  atau  $\alpha = 0,05$ . Daerah kritik yang dipilih adalah daerah yang memenuhi  $P\{\text{kesalahan tipe I}\} \leq \alpha$  sedemikian hingga *power* dibawah  $H_1$  maksimum (Somayasa, 2019).

Perbandingan fungsi *power* dari uji Kolmogorov-Smirnov dan beberapa uji goodness of fit yang lain seperti uji Cramer-Von Mises dan uji Anderson-Darling dapat dilihat dari nilai *power* masing-masing uji. Semakin tinggi nilai *power* yang diperoleh suatu uji, maka semakin baik pula uji

tersebut dalam hal membuat keputusan dari suatu hipotesis uji.

Karena nilai *power* tidak dapat dihitung secara analitik, maka dalam kasus ini *power* dari uji Kolmogorov-Smirnov dapat didekati dengan simulasi Monte Carlo dengan bantuan program R. Untuk menghitung *power* dari uji Kolmogorov-Smirnov, uji Cramer-Von Mises, dan uji Anderson-Darling untuk berbagai macam ukuran sampel yang mengikuti model distribusi di bawah  $H_1$  dapat dilakukan sesuai algoritma simulasi sebagai berikut:

1. Menentukan ukuran-ukuran sampel yang akan digunakan sebagai bahan perbandingan.
2. Menentukan taraf signifikansi  $\alpha$  yang akan digunakan.
3. Membangkitkan sampel acak  $X_1, X_2, \dots, X_n$  dari distribusi dibawah  $H_1$ .
4. Menghitung statistik  $D_n, \omega_n^2$ , dan  $A_n^2$ .
5. Langkah 3 dan 4 diulangi sebanyak  $M$  kali.
6. Menghitung nilai pendekatan dari fungsi *power* dengan rumus:

- a. Untuk statisti kuji Kolmogorov-Smirnov

$$power = \frac{1}{M} |\{D_n > k_{n;1-\alpha}\}|$$

- b. Untuk statistik Cramer Von-Mises

$$power = \frac{1}{M} |\{\omega_n^2 > k_{n;1-\alpha}\}|$$

- c. Untuk statistik Anderson-Darling

$$power = \frac{1}{M} |\{A_n^2 > k_{n;1-\alpha}\}|$$

7. Membandingkan nilai *power* dari masing-masing uji tersebut.

Hasil simulasi terhadap *power* di bawah  $H_1$  dibawah dengan bantuan program R ditampilkan pada Tabel 3.1 dan Tabel 3.2. Simulasi pertama dilakukan dengan membangkitkan data yang mengikuti distribusi di bawah  $H_1$  yang diasumsikan sebagai  $u \sim EXP(\lambda)$ . Selanjutnya, simulasi berikutnya dilakukan dengan mengikuti distribusi yang diasumsikan sebagai  $u \sim WEI(\theta, \beta)$ . Simulasi ini dilakukan dengan menentukan berbagai nilai  $n$  dan  $M$ . Taraf signifikansi  $\alpha$  yang digunakan adalah sebesar 1% dan 5%.

Angka-angka pada Tabel 3.1 dan 3.2, menunjukkan bahwa untuk uji Kolmogorov-Smirnov, uji Cramer Von-Mises, dan uji Anderson-Darling, nilai *power* yang diperoleh dari sampel yang dibangkitkan dari distribusi eksponensial dan distribusi Weibull untuk berbagai  $n$  dan  $M$  dengan  $\alpha = 1\%$  dan  $\alpha = 5\%$  adalah 1. Jadi dengan nilai *power* yang maksimal tersebut mengidentifikasi bahwa ketiga uji mempunyai kemampuan yang tinggi untuk mendeteksi  $H_0$  yang salah.

### 3.3 Aplikasi pada data daya tahan marmut

Pada bagian ini, akan dilakukan uji *Goodness of Fit* dengan menggunakan statistik uji Kolmogorov-Smirnov dan yang dikaji pada sub Bab 3.1 dengan

tujuan untuk apakah data yang digunakan berasal dari populasi yang berdistribusi gamma terboboti. Data yang digunakan dalam penelitian ini adalah waktu (dalam hari) kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen. Data ini bersumber dari makalah, dengan judul “*Goodness of Fit Tests and Power Comparisons For Weighted Gamma Distribution*” oleh Sangla, Neetu, Kanchan Jain, dan Suresh Kumar Sharma.

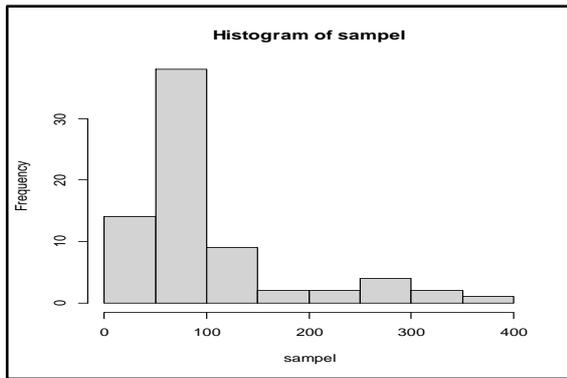
**Tabel 3.1.** Nilai *power* untuk  $u \sim \exp(\lambda)$

No.	n	B	$\lambda$	Power					
				$\alpha = 1\%$			$\alpha = 5\%$		
				KS	CVM	AD	KS	CVM	AD
1	50	500	3	1	1	1	1	1	1
2	100	1000	3	1	1	1	1	1	1
3	150	2500	3	1	1	1	1	1	1
4	300	5000	3	1	1	1	1	1	1
5	500	7500	3	1	1	1	1	1	1
6	800	9000	3	1	1	1	1	1	1
7	1000	10000	3	1	1	1	1	1	1

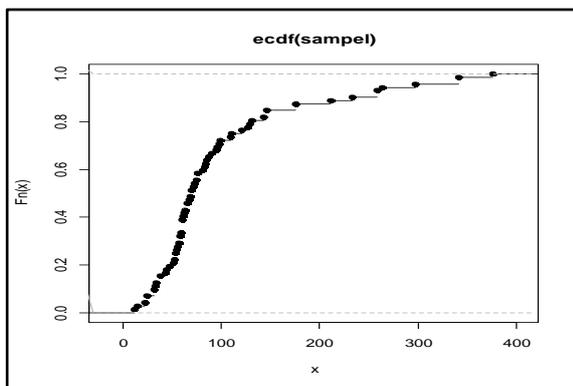
**Tabel 3.2.** Nilai *power*  $u \sim WEI(\theta, \beta)$

No.	n	B	$(\theta, \beta)$	Power					
				$\alpha = 1\%$			$\alpha = 5\%$		
				KS	CVM	AD	KS	CVM	AD
1	50	500	(3,5)	1	1	1	1	1	1
2	100	1000	(3,5)	1	1	1	1	1	1
3	150	2500	(3,5)	1	1	1	1	1	1
4	300	5000	(3,5)	1	1	1	1	1	1
5	500	7500	(3,5)	1	1	1	1	1	1
6	800	9000	(3,5)	1	1	1	1	1	1
7	1000	10000	(3,5)	1	1	1	1	1	1

Berdasarkan informasi dari data, waktu kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen mencapai nilai minimum sebesar 12 hari, nilai maksimum sebesar 376 hari, nilai median sebesar 70 hari, dan nilai mean sebesar 99,82 hari. Secara visual, ditampilkan histogram dan grafik fungsi distribusi empiris dari data yang digunakan. Kedua grafik ini dibangun dengan bantuan Program R yang ditampilkan pada Gambar 3.3 dan Gambar 3.4. Untuk lebih meyakinkan bahwa data tersebut berdistribusi gamma terboboti, maka akan dilakukan uji *Goodness of Fit* menggunakan statistik Kolmogorov-Smirnov.



Gambar 3.3. Histogram waktu hidup marmut.



Gambar 3.4 CDF empiris waktu hidup marmut.

Berdasarkan prosedur uji menggunakan statistik Kolmogorov-Smirnov, keputusan menolak atau menerima  $H_0$  pada tingkat signifikansi  $\alpha$  diambil dengan melihat kriteria berikut:

- a. Jika  $D_n \geq k_{n;1-\alpha}$ , maka  $H_0$  ditolak dan disimpulkan data tidak berasal dari distribusi gamma terboboti.
- b. Jika  $D_n < k_{n;1-\alpha}$ , maka  $H_0$  diterima dan disimpulkan data berasal dari distribusi gamma terboboti.

dimana  $k_{n;1-\alpha}$  merupakan kuantil  $1 - \alpha$  dari distribusi peluang  $D_n$ . Kuantil-kuantil ini diaproksimasi dengan menggunakan simulasi Monte Carlo dengan bantuan *software* R.

Dari hasil perhitungan statistik Kolmogorov-Smirnov pada data waktu kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen menghasilkan  $D_n = 0.6268$  seperti yang terlampir pada Lampiran 10. Sedangkan untuk nilai  $k_{n;1-\alpha}$  untuk taraf signifikansi  $\alpha = 5\%$  menghasilkan  $k_{72;0,95} = 1,4904$ , maka  $H_0$  diterima sebab  $D_{72} < k_{72;0,95}$  yaitu  $0.6268 < 1,4904$ . Sehingga dapat disimpulkan data waktu kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen berasal dari populasi berdistribusi gamma terboboti. Inferensi yang diasumsikan hanya untuk distribusi gamma terboboti, tidak dapat digunakan untuk distribusi yang lain.

Selanjutnya, hasil pengujian ini akan disinkronkan dengan uji goodness of fit berdasarkan *ECDF* yang lain yaitu Cramer-von Moses. Hasil perhitungan statistik Cramer-von Mises pada data waktu kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen menghasilkan  $W_{72}^2$  sebesar 0,1782. Berdasarkan tabel nilai kritis statistik Cramer-von Mises untuk  $\alpha = 5\%$  diperoleh  $W^2$  sebesar 0,6568, maka  $H_0$  diterima sebab nilai  $W_{72}^2 < 0,6568$ . Sehingga dapat disimpulkan data waktu kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen berasal dari populasi berdistribusi gamma terboboti.

Dari hasil uji tersebut, uji Kolmogorov-Smirnov dengan menggunakan distribusi gamma terboboti menghasilkan hasil uji yang sinkron terhadap hasil uji statistic Cramer-von Mises.

#### 4. Penutup

##### 4.1 Kesimpulan

Berdasarkan pembahasan pada bagian 3, hal-hal yang dapat disimpulkan adalah sebagai berikut:

1. Proses empirik  $\sqrt{n}(\hat{F}_n(t) - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  konvergensi dalam distribusi ke  $B^0(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$  yaitu suatu proses Gaussian dengan mean 0 dan variansi  $F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda})(1 - F_0(t_k; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))$ ,  $\forall t \in \mathcal{R}$ .  
Jadi  $B^0(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda})) \sim N(0, F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda})(1 - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda})))$ . Sebagai akibat dari teorema pemetaan kontinu, maka  $\sqrt{n} \sup_{0 \leq t \leq 1} |\hat{F}_n(t) - F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda})|$  konvergen dalam distribusi ke  $\sup_{0 \leq t \leq 1} |B^0(F_0(t; \hat{\alpha}, \hat{\beta}, \hat{\lambda}))|$  untuk  $n \rightarrow \infty$ .
2. Pada penelitian ini, apabila data dibangkitkan dari distribusi Weibull dan distribusi Exponensial fungsi power yang diperoleh untuk uji Anderson Darling, uji Kolmogorov-Smirnov, dan uji Cramer-von Mises untuk berbagai ukuran ( $n$ ) dan ( $M$ ) dengan  $\alpha = 1\%$  dan  $\alpha = 5\%$  yaitu 1. Jadi, ketiga uji mempunyai power yang sama.
3. Aplikasi uji Kolmogorov-Smirnov pada data waktu kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen menerima  $H_0$  pada  $\alpha = 5\%$ , karena  $D_{72} < k_{72;0,95}$  yaitu  $0.8871 < 1,4904$ . Sedemikian sehingga, hasil pengujian data waktu kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen menunjukkan bahwa data berasal dari distribusi gamma terboboti.

##### 4.2 Saran

Dalam penelitian ini penulis menunjukkan uji Kolmogorov-Smirnov terhadap distribusi gamma terboboti dengan mengaplikasikan pada data waktu

kelangsungan hidup marmut yang terinfeksi basil tuberkulum virulen. Pada penelitian selanjutnya penulis menyarankan agar peneliti menunjukkan secara analitik dengan menghitung kuartil-kuantil dari statistik uji Kolmogorov-Smirnov untuk ukuran sampel tak hingga.

**Ucapan Terimakasih.** Penulis menyampaikan terimakasih yang setulus-tulusnya kepada dosen pembimbing atas segala curahan perhatiannya sehingga tulisan ini dapat diselesaikan dengan baik. Penulis juga menyampaikan terimakasih kepada seluruh pihak yang turut andil dalam penelitian ini baik langsung maupun tidak langsung.

#### Daftar Pustaka

- Ahmad, A., dan Ahmad, S. P. (2019). Weighted analogue of inverse Lévy distribution: Statistical properties and estimation. *Journal of Applied Probability and Statistics*, 14(1), 25–37.
- Billingsley, P. 1995. *Probability and Measure (Third Edition)*, John Wiley & Sons, inc. New York.
- Daniel, W. (1989). *Statistika Nonparametrik Terapan*. (A. T. Kantjono, Penerj.) Jakarta: PT. Gramedia.
- Diantoro, A. 2013. *Model Statistik untuk Data Pencemaran Udara oleh Ozon Ekstrem di Pekanbaru*. Pekanbaru: Universitas Islam Negeri Sultan Syarif Kasim Riau.
- Fallo, J. O. Setiawan, A. dan Susanti, B. (2013). *Uji Normalitas Berdasarkan Metode Anderson-Darling, Cramer-Von Mises dan Lilliefors Menggunakan Metode Bootstrap Prosiding*. ISBN: 978-979-16353-9-4. Universitas Kristen Satya Wacana. Salatiga.
- Lehmann, E. L dan Romano, J. P. (2005). *Testing Statistical Hypotesis (Third Edition)*, Springer Science+Business Media, Inc., New York.
- Ross, S.M., 1996, *Stochastic Processes (Second Edition)*, John Wiley & Sons, inc., New York.
- Saghir, A., Hamedani, G. G., Tazeem, S., dan Khadim, A. (2017). Weighted Distributions: A Brief Review, Perspective and Characterizations. *International Journal of Statistics and Probability*, 6(3), 109.
- Sen, S., Chandra, N., dan Maiti, S. S. (2017). the Weighted Xgamma Distribution: *Reliability and Statistical Studies*, 10(1), 43–58.
- Seruling, R. J., 1980. *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, inc. New York.
- Singla, N., Jain, K., dan Sharma, S. K. (2016). Goodness of fit tests and power comparisons for Weighted Gamma distribution. *Revstat Statistical Journal*, 14(1), 29–48.
- Somayasa, W. 2019. *Statistika Matematika (Teori Peluang dan Dasar-Dasar Inferensi Parameter)*. Yogyakarta: Deepublish Budi Utama.
- Suyanto. (2006), Uji Goodness of Fit Distribusi Eksponensial Tergeneralisir Dengan Metode Anderson-Darling, *Skripsi*, Universitas Airlangga, Surabaya.
- Ahmad, A., & Ahmad, S. P. (2019). Weighted analogue of inverse Lévy distribution: Statistical properties and estimation. *Journal of Applied Probability and Statistics*, 14(1), 25–37.
- Warella, R. Y., Wattimanela, H. J., & Ilwaru, V. Y. I. (2021). Sifat-Sifat Dan Kejadian Khusus Distribusi Gamma. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 15(1), 047–058.

Diterima pada tanggal 26 Mei 2022.  
Terbit online pada tanggal 28 Juli 2022